

# Processing and Probability Analysis of Pulsed Terahertz NDE of Corrosion Under Shuttle Tile Data

Robert F. Anastasi<sup>1</sup>, Eric I. Madaras<sup>2</sup>, Jeffrey P. Seebo<sup>3</sup>, and Thomas M. Ely<sup>3</sup>

<sup>1</sup>U.S. Army Research Laboratory, Vehicle Technology Directorate, AMSRD-ARL-VT-MD, Nondestructive Evaluation Sciences Branch, NASA Langley Research Center, MS231, Hampton, VA 23681

<sup>2</sup>NASA Langley Research Center, Nondestructive Evaluation Sciences Branch, MS231 Hampton, VA 23681

<sup>3</sup>Lockheed Martin, NASA Langley Research Center, MS231, Hampton, VA 23681

## ABSTRACT

This paper examines data processing and probability analysis of pulsed terahertz NDE scans of corrosion defects under a Shuttle tile. Pulsed terahertz data collected from an aluminum plate with fabricated corrosion defects and covered with a Shuttle tile is presented. The corrosion defects imaged were fabricated by electrochemically etching areas of various diameter and depth in the plate. In this work, the aluminum plate echo signal is located in the terahertz time-of-flight data and a threshold is applied to produce a binary image of sample features. Feature location and area are examined and identified as corrosion through comparison with the known defect layout. The results are tabulated with hit, miss, or false call information for a probability of detection analysis that is used to identify an optimal processing threshold.

**Keywords:** Terahertz, Nondestructive Evaluation, Corrosion, Probability

## 1. INTRODUCTION

Shuttle tiles are inspected for damage and periodically removed so underlying structure can be examined. Corrosion is a basic structural degradation that can occur due to the Shuttle's unique operating environment. For aluminum alloys, corrosion generally develops as pitting or thinning, and in general changes a nominally smooth surface to an uneven and irregular surface, which can then result in cracking and further structural damage. A potential method for detecting corrosion under tiles is pulsed Terahertz (THz) nondestructive evaluation (NDE).

Pulsed THz NDE technology offers a non-contact and high-resolution inspection method. Its general frequency regime is between 300GHz and 3THz and it has a free space wavelength range of 1mm to 0.1mm. This is a region of the electromagnetic spectrum between the microwave and infrared bands. This method has found NDE application for aerospace material NDE [1, 2]. The THz electromagnetic radiation can penetrate the nonconductive Shuttle tiles and inspect the underlying metallic surface. Inspecting for corrosion under tiles with pulsed THz NDE has been demonstrated and the measurements indicated that corroded regions with depths greater than 0.13mm could be detected through a one-inch tile [3]. In this past work, evaluations were based on visual inspection of the THz image where dark areas or spots that stood out from the background were equated to corrosion. This was adequate for the evaluations at that time, but could be biased by the person performing the evaluation.

In this paper an automated data processing method is developed and applied to pulse THz NDE scans of corrosion under a Shuttle tile to help remove bias in data evaluation and a probability of detection (POD) analysis is performed to quantify the evaluation.

## 2. MEASUREMENTS

The pulsed THz system used in this work has been described in detail elsewhere [4] and was initially designed to inspect up to 30cm of Sprayed-On Foam Insulation on the External Tank of the Space Shuttle. The transceiver for this system, pictured and schematically shown in Figure 1, is a transmitter and receiver in a confocal arrangement. This confocal arrangement uses a 15cm diameter, 285mm focal length lens to concentrate the THz energy on the sample to a spot diameter of approximately 3mm. Bandwidth at half-height is  $\sim 100\text{GHz}$  to  $500\text{GHz}$  and its full width is  $\sim 30\text{GHz}$  to  $1.0\text{THz}$ .

The system was integrated with a data acquisition and control system that translated the THz transceiver along two dimensions and acquires and stores temporal waveforms at incremental scan points. The waveforms are composed of 2048 points with a duration of 320ps, which corresponds to a propagation path length in air of 96mm.

## 3. SAMPLE AND THZ IMAGE

The sample used in this work was a test article consisting of an array of tiles 24.5mm thick attached to an aluminum substrate. This panel, originally used in an impact study, was modified by removing a tile ('tile 745'), electrochemically etching corroded regions in the aluminum substrate [4], and reattaching the tile in the manner used on the Shuttle.. General tile properties in the THz regime were characterized and THz scans of this test article were conducted [4]. One of these scans was evaluated in this study. Tile properties were characterized by measuring reflections from a metal substrate with and without an intermediate tile. Thick, medium, and thin tiles used in this characterization were approximately 15cm square and 76mm, 48mm, and 12mm thick respectively. From the measurements, signal attenuation was found to be 3.06dB/cm at a peak frequency of 250GHz, the average wave velocity was  $2.78 \pm 0.04 \times 10^8$  m/s, and the real part of the refractive index for the material was  $1.08 \pm 0.02$ .

A schematic of the test article, a picture of the corroded regions and the resulting THz amplitude image of 'tile-745' are shown in Figure 2. Figure 2a shows the arrangement of the tiles on the panel. The number on the tile corresponds to its identification number and the 'D' labels indicated a tile, with impact damage. Tile-745 was scanned with the pulsed THz system in a raster scan pattern. The scan was 151x151 points with x- and y-axis resolution of 0.1mm. Figure 2b shows a picture of the corrosion array under tile 745, and Figure 2c shows the resulting THz image of tile 745. This THz image is a signal amplitude image generated by gating the THz signal substrate echo and plotting its normalized amplitude as a gray scale value. In this image low amplitude values are black or dark gray while large amplitude values display as light gray or white.

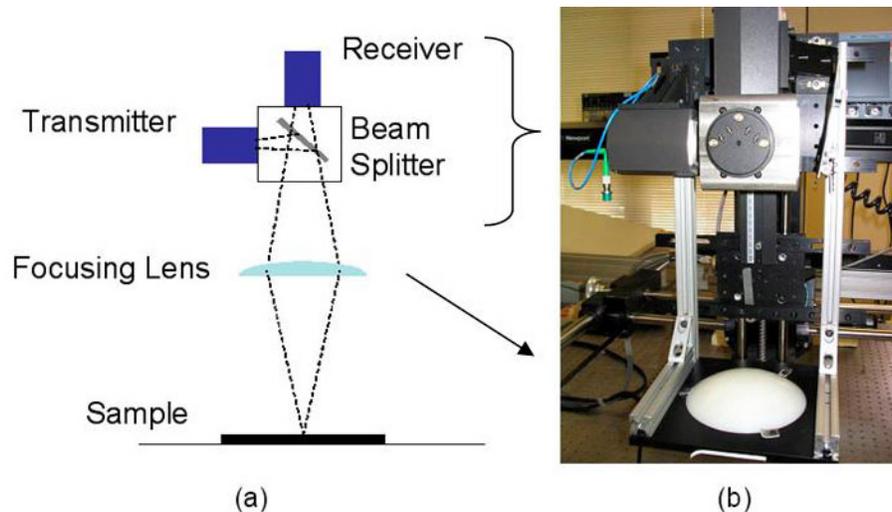


Figure 1. Transceiver confocal arrangement (a) schematic and (b) photograph.

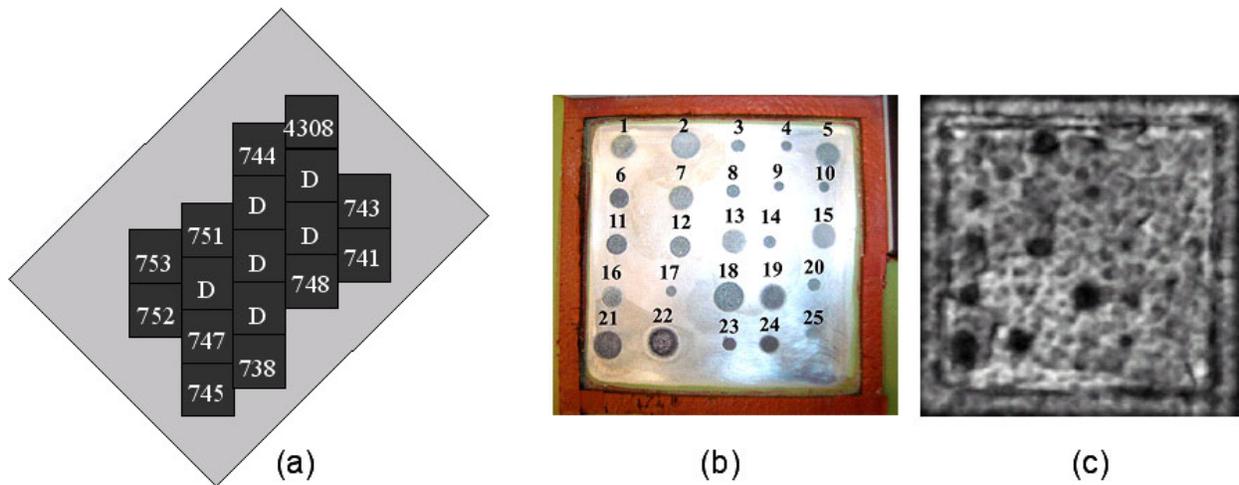


Figure 2. Test article (a) schematic of the panel showing tile arrangement, (b) a picture of the corrosion regions under tile 745 with spots numbered, and (c) its THz image through the attached tile

The overall gray level variation appears mottled due to general scattering of THz radiation as it travels through the tile. Corroded regions appear darker because they scatter the incident radiation more than the unaltered substrate regions. The image includes effects due to tile defects, interfaces, and substrate damage averaged over the focused path of the THz energy. In this work, THz imaging of these features were not individually classified. Classifying spots in the THz image as either tile or substrate defects is a project currently in progress and is not reported here.

#### 4. DATA IMAGE PROCESSING

A THz processing program was developed to extract corrosion spot data and eliminate bias introduced by the person inspecting the sample. This program operates on the scan data to produce a binary image of corrosion sample features. It was found that conventional image processing techniques such as edge detection techniques were too sensitive, detecting corrosion edges as well as edges from mottled areas in the image. To reduce sensitivity, smoothing was applied, but then detail was lost. The following procedure offered more control over the final image using row and column slope, a threshold level, and filtering to generate the desired corrosion spot image.

In the first step, time domain signals in the scan data were reduced from 2048 points to a single value that represented the peak-to-peak amplitude of the reflected signal returned from the tile/aluminum interface for every measured spatial location. This process replicated the gray scale image shown in Figure 2c and is referred to as the raw image. In the second step, a binary image was generated by applying a Gaussian low pass spatial filter to the raw image, normalizing and scaling the result, and then a threshold level. This process reduced small speckle noise and smoothed edges of spatially larger spots. In the third step, two images were generated. One image was based on row (a horizontal line in the raw image) slope values and the other on column (a vertical line in the raw image) slope values. These two slope images and the previous image were added and then normalized and rescaled. The result was an enhanced binary image. In the final fifth step, the enhanced binary image was passed to a spot analysis section of the software where spots were counted and area and center coordinates calculated and tabulated. Example output for two processing threshold levels is shown in Figure 3. These images show decreasing spot detection for increasing threshold level.

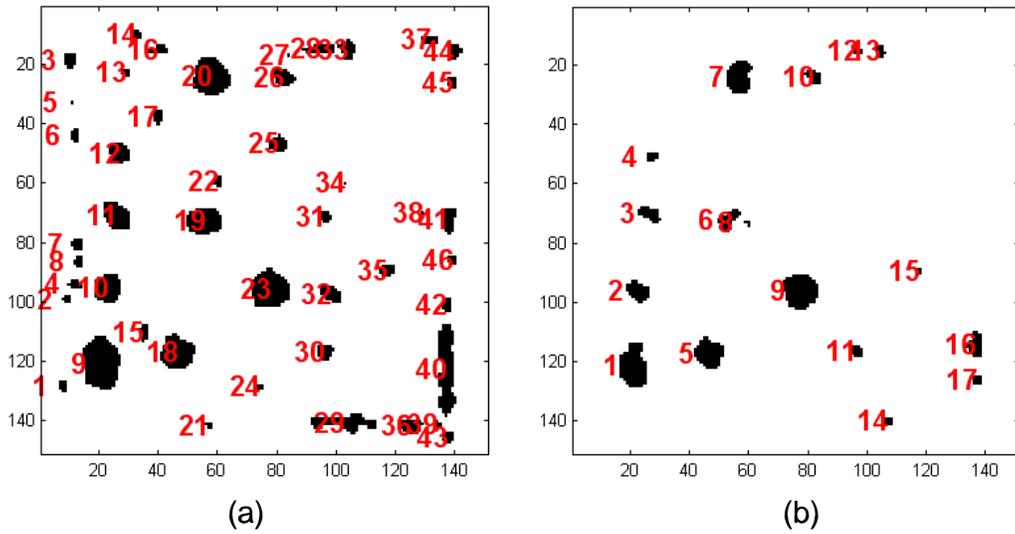


Figure 3. Corrosion features for threshold level (a) 1.20 and (b) 1.50

Features in these images were examined and identified as corrosion through comparison with the known corrosion defect map and hit, miss, and false calls were determined. A hit is defined as a correctly identified corrosion spot, a miss is an undetected spot, and a false call is a detected spot that is not corrosion. Table I shows a tabulation of hit, miss, and false call values for a series of threshold levels. The number of false calls is very high for the low threshold level compared to the other levels.

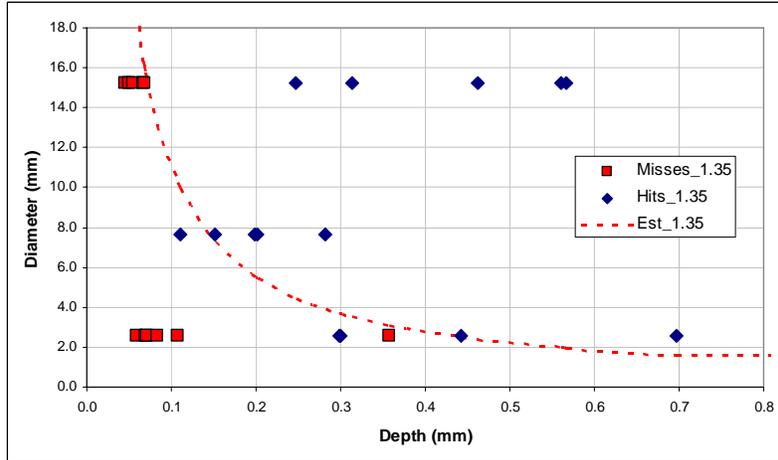
A detection threshold limit was estimated from the processing results by plotting hit/miss data against corrosion spot diameter and depth and fitting the boundary between hits and misses with an appropriate function where the function in this case was the depth raised to a power and scaled. Hit/miss plots against diameter and depth are illustrated in Figure 4. An estimated detection curve for a processing threshold of 1.20 is shown in Figure 4a and a family of curves for five processing thresholds is shown in Figure 4b. These results illustrate how processing threshold influences measurement outcome and that bias by the inspection personnel has been replaced with the bias of choosing a threshold level in the data processing.

Table I. Hit/Miss and False Calls for a series of Threshold levels

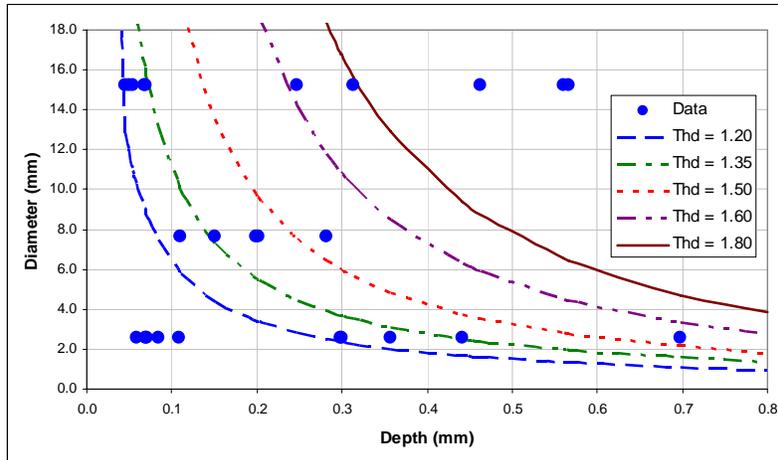
Processing Threshold Level	0.90	1.00	1.10	1.20	1.30	1.40	1.50	1.60	1.70	1.80
Hits	18	16	16	16	15	14	11	7	4	3
Misses	7	9	9	9	10	11	14	18	21	22
False Calls	44	38	36	60	25	10	5	1	1	0

## 5. POD ANALYSIS

To quantify the processing results as a function of threshold a POD analysis was conducted. An analysis program was written to calculate the POD of hit/miss data based on the maximum likelihood estimate (MLE) of the log-logistic distribution. This distribution was one of seven examined by Berens and Hovey [5], they concluded that the log-logistic distribution gave the most consistent results for POD crack analysis. In the analysis presented here ‘crack’ was replaced with the defect parameter of corrosion depth or diameter.



(a)



(b)

Figure 4. Estimate detection limits versus threshold level (a) for threshold level of 1.35 and (b) for threshold levels of 1.20, 1.35, 1.50, 1.60, and 1.80.

The functional form of the log-logistic distribution is: [5-7]

$$P_i = \frac{\exp(\alpha + \beta \ln(a_i))}{1 + \exp(\alpha + \beta \ln(a_i))} \quad (1)$$

where  $P_i$  is the probability of detecting flaw  $i$ ,  $a_i$  is the size of flaw  $i$ , and  $\alpha$  and  $\beta$  are location and scale parameters defining the distribution. The analysis program was verified by reproducing published POD curves for crack data [7].

The analysis starts by grouping the hit/miss data into intervals. For each interval a proportion of hits (POH) was determined as the ratio of defects detected to the total number of defects in that interval and resulted in a set of data; midpoint of the interval and POH. A linear regression of this data set data gives starting values used in applying the MLE of the log-logistic distribution. The resulting MLE estimates  $\hat{\alpha}$  and  $\hat{\beta}$  substituted into equation [1] give the mean POD curve. The 95% POD lower confidence level (LCL) curve is calculated in a manner similar to the POD curve with

1.96 times the standard error subtracted from the exponential terms in equation [1], where 1.96 is the 97.5 percentile point of the normal distribution [7, 8].

POD curves for the 1.20 processing threshold using defect parameters corrosion depth and diameter are shown in Figure 5. These plots show the hit/miss data plotted as a '1' or a '0', POH data, mean POD curve, and the 95% POD LCL curve. The POD curve in Figure 5a shows the 90% mean POD limit on corrosion depth was 0.3 mm and the 95% LCL does not exceed 90% probability (the 90/95 POD was not reached) over the range of corrosion depth. Figure 5b shows the 90% mean POD was not reached over the range of corrosion diameters and the 95% lower confidence level does not exceed 90% probability (90/95 POD was not reached) over the range of diameter. These plots show the POD of depth is greater than POD of diameter and thus may be a better parameter for corrosion evaluation

To determine the optimal processing threshold a series of mean POD curves versus thresholds were compared and are shown in Figure 6. The POD curves group into regions of high, medium, and low level probability. The high probability level includes thresholds of 0.9 to 1.4 and medium and low probability levels include ranges from 1.5 to 1.6 and 1.7 to 1.8, respectively. If only these curves were considered then one would choose a threshold level that offers the highest probability, but these high levels have a large number of false calls. The number of false calls that can be tolerated needs to be determined by engineering and cost analysis. In the literature, a false calls rate of 5% is suggested to ensure true POD representation [9, 10] where false call rate is defined as the number of false calls divided by the total number of observations. Using the data in Table I the false call rate as a function of processing threshold was calculated and is shown in Figure 7. From this plot the suggested 5% false call rate shown to be near the 1.6 threshold level. The number of misses at this point exceeds the number of hits by more than a factor of two and suggests that the 5% rate is too conservative for this analysis. The series of POD curves in Figure 6 may be a means of justifying a less conservative false call rate and resulting optimum processing threshold level. Since the 1.5 and 1.6 levels group together, they can be considered a lower and upper bound on acceptable false calls. At this 1.5 level, Figure 7 shows the false call rate to be between 15% and 20%.

Using the 1.5 processing threshold and the estimated detection limit, a high confidence detection region was estimated and is shown in Figure 8. This region highlights the corrosion depth and diameter parameters that would be detected with high confidence.

Since the 90/95 POD was not reached, a beta version design of experiments software package [11] was used to examine some of the data and provide recommendations for validating the POD. The program uses observed hit/miss measurements and a moving average to dynamically group defect parameters into class widths that results in a 'point estimation of a hit.' The software gave similar trends to the analysis described above verifying the 90/95 POD was not reached. It recommended adding 26 samples at a depth of 0.69mm, adding 26 samples at 0.56mm depth, and adding 29 samples with the corrosion depths of 1.39mm or greater. Thus for POD validation a larger data set was recommended.

## 6. SUMMARY

This paper examined data processing and probability analysis of pulsed terahertz NDE scans of corrosion defects under a Shuttle tile. A data processing program was written to extract corrosion spot data and eliminate bias posed by the person inspecting the sample. Procedures developed use row and column slope, a threshold level, and filtering to generate a binary image of corrosion sample features which were then compared to the original corroded region map to identify hit, miss, and false calls. This processing transferred bias from individual observations to that of choosing a threshold level. To remove this bias and identify an optimal threshold level a POD analysis was conducted that took into account the misses and hits identified. A series of POD curves were generated over a range of threshold levels and when plotted together they divided into high, medium, and low probability groups. This processing suggested an optimal processing threshold that minimized false calls without eliminating all but, the largest/deepest corroded regions. Using this identified threshold, a high confidence detection region was estimated by combining POD curves and estimated detection limit. This detection region gives the corrosion depth and diameters that would most likely be detected for the given problem parameters. A design of experiments software recommended a larger data set to validate POD results and showed the same trends as this analysis which provided a means of optimizing the processing level that discriminates between corroded regions and noise for a given measurement process and tile thickness parameter.



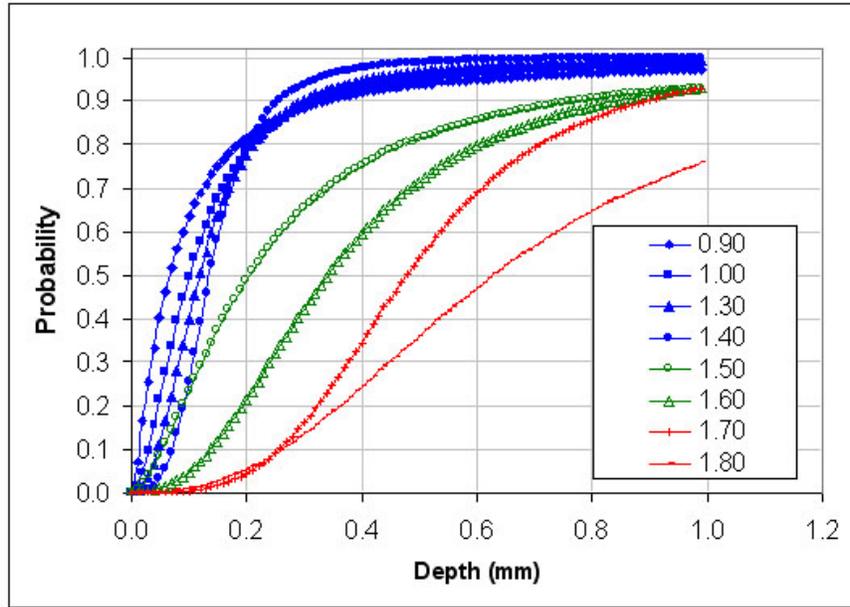


Figure 6. POD curves for a series of processing thresholds divided into groups. The high probability level includes thresholds of 0.9 to 1.4. The medium and low probability levels include ranges from 1.5 to 1.6 and 1.7 to 1.8, respectively

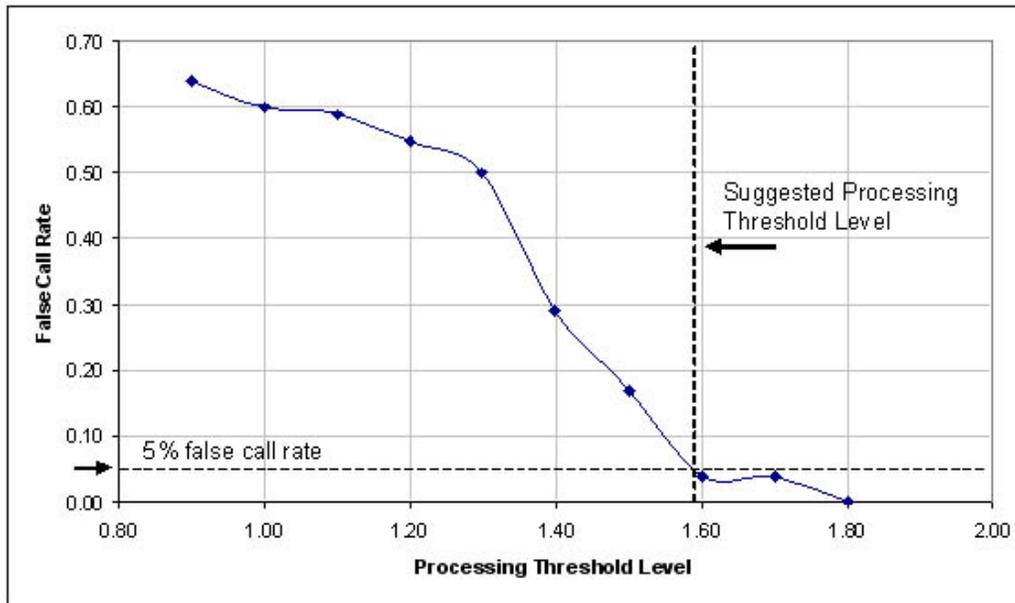


Figure 7. False call rate as a function of processing threshold. The literature suggested false call rate of 5% gives a threshold processing level of ~1.60.

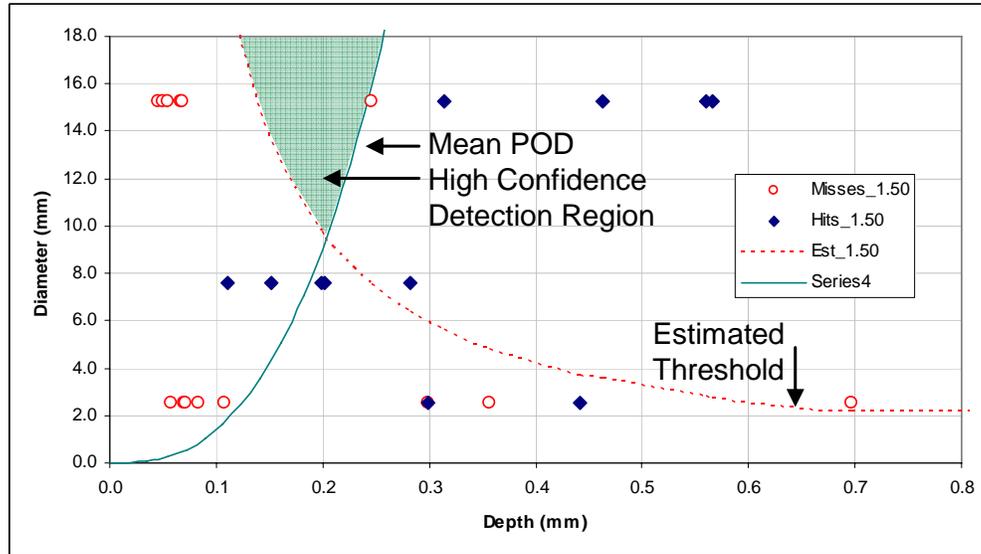


Figure 8. Estimate High Confidence Detection Region for the processing threshold of 1.5

## REFERENCES

- [1] Anastasi, R.F., Madaras, E.I., Seebo, J.P. and Winfree, W.P., "Terahertz NDE for Aerospace Applications," published in "Ultrasonic and Advanced Methods for Nondestructive Testing and Material Characterization," Edited by C.H. Chen, published by World Scientific Publishing, June (2007)
- [2] Zimdars, D, White, J.S, Stuk, G., and Fichter, G., "Large Area Time Domain Terahertz (T-Ray) Imaging Non-Destructive Evaluation for Security and Aerospace," published in "Ultrasonic and Advanced Methods for Nondestructive Testing and Material Characterization," Edited by C.H. Chen, published by World Scientific Publishing, June (2007)
- [3] Anastasi, R.F., Madaras, E.I., Seebo, J.P., Smith, S.W., Lomness, J.K., Hintze, P.E., Kammerer, C.C., Winfree, W/P., and Russell, R.W., "Terahertz NDE Application for Corrosion Detection and Evaluation Under Shuttle Tiles," SPIEs 14th Annual International Symposium on Smart Structures and Materials and Nondestructive Evaluation and Health Monitoring, San Diego, CA, March 18-22, (2007)
- [4] Zimdars, D., Valdmanis, J. A., White, J. S., Stuk, G., Williamson, S., Winfree, W.P, and Madaras, E. I., "Technology and Applications of Terahertz Imaging Non-Destructive Examination: Inspection of Space Shuttle Sprayed On Foam Insulation," Review of Progress in Quantitative Nondestructive Evaluation, Golden, Colorado, 25-30 July (2004)
- [5] Berens, A.P. and Hovey, P.W., "Characterization of NDE Reliability," Review of Progress in Quantitative NDE, Vol. 1, eds Thompson and Chimenti, Plenum Press, New York, pp579-585, (1982)
- [6] Forsyth, D.S. and Fahr, A., "An Evaluation of Probability of Detection Statistics," RTO-MP-10, AC/323(AVT)TP/2, RTO Meetings Proceedings 10, Airframe Inspection Reliability under Field/Depot Conditions, NATO Research and Technology Organization, France, November, (1998)
- [7] Bullock, M. Forsyth, D. and Fahr, A., "Statistical Functions and Computational Procedures for the POD Analysis of Hit/Miss NDI Data," LTR-ST-1964, National Research Council Canada, (1994)
- [8] Wikipedia contributors. 1.96. Wikipedia, The Free Encyclopedia. June 6, 2008, 15:12 UTC. Available at: <http://en.wikipedia.org/w/index.php?title=1.96&oldid=217547884>. Accessed October 17, (2008)
- [9] Petrin, C.L., Vukelich, S.L., and Annis, C.A., "A Recommended Methodology for Quantifying NDE/NDI Based On Aircraft Engine Experience," Advisory Group for Aerospace Research and Development (AGARD), AGARD-LS-90, April (1993)
- [10] Simsir, M. and Ankara, A., "Comparison of Two Non-Destructive Inspection Techniques on the Basis of Sensitivity and Reliability," Materials and Design 28, pp 1433-1439, (2007)
- [11] Generazio, E.R., "Directed Design of Experiments for Validating Probability of Detection Capability of NDE Systems (DOEPOD)," Review of Progress in Quantitative NDE: 34th Annual Review of Progress in Quantitative Nondestructive Evaluation. AIP Conference Proceedings, Volume 975, pp. 1693-1700 (2008)